UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/282,619	03/31/1999	QUAN G. CUNG	AT9-99-037	8855

7590 07/15/2002

ANDREW J DILLON
FELSMAN BRADLEY GUNTER & DILLON
LAKEWOOD ON THE PARK SUITE 350
7600 E NORTH CAPITAL OF TEXAS HIGHWAY
AUSTIN, TX 78731

EXAMINER	
DAY, HERNG-DER	
ART UNIT	**PAPER NUMBER**
2123	

DATE MAILED: 07/15/2002

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	09/282,619	CUNG ET AL.
	Examiner	Art Unit
	Herng-der Day	2123

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on 29 May 2002 .

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-25 is/are pending in the application.

4a) Of the above claim(s) 2 and 7-12 is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1, 3-6, 13-25 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 31 March 1999 is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

1. This communication is in response to Applicant's Reply to Office Action Dated April 23, 2002 mailed on May 29, 2002. Claims 1, 3-6, 13-14, and 20-25 were amended; claims 2 and 7-12 were cancelled; claims 1, 3-6, and 13-25 are pending.

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claims 1, 3-6, and 13-25, are rejected under 35 U.S.C. 103(a) as being unpatentable over Piatetsky-Shapiro, "Discovery, Analysis, and Presentation of Strong Rules", in "Knowledge Discovery in Database", AAAI/MIT Press, 1991, in view of Simoudis et al., U.S. Patent No. 5,692,107 issued on November 25, 1997, and further in view of Dash et al., "Dimensionality Reduction of Unsupervised Data", Proceedings, Ninth IEEE International Conference on Tools with Artificial Intelligence, Nov. 1997.

3.1 Regarding claims 1 and 3-6, Piatetsky-Shapiro teaches comparing, see section 13.3, Rule-Interest Measures, attribute values for samples having one or more desired attributes and respective values, see section 13.5, KID3 Algorithm, to attribute values for all samples, see section 13.4, Precomputing Field-Value Statistics (claim 1). Therefore, patterns of full data set can be estimated by sample-derived rules with an estimated accuracy, see Abstract, i.e., a

predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. However, Piatetsky-Shapiro does not teach the selection of a subset of available attributes.

Simoudis et al. teaches the selection of a data analysis module, for example, a statistical module, to perform data mining to the selected data set (claim 1). The user sets module-specific parameters. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results. See col. 4, line 42 through line 57. However, Simoudis et al. does not teach the selection of a subset of available attribute based on entropy measure.

Dash et al. teaches an entropy measure (claim 3) for determining the relative importance of variables; see section 2. Dash et al. also discloses a simple way to decide how many variables should be kept for a task by choosing the first d variables if it is known that an application only needs d variables (claim 4). See section 3. Thus the user gains insight into the data after the important original features are known.

In order to generate a predictive model for sample population having one or more desired attributes and respective values one of ordinary skill in the art would first be motivated by Piatetsky-Shapiro to compare attribute values for samples having one or more desired attributes and respective values to attribute values for all samples (claim 1). Then to determine the statistical difference in a scientific and more efficient way one of ordinary skill in the art would be motivated by Dash et al. to apply an entropy measure (claim 3) to the above-mentioned data sets (claim 1) for data mining. Data mining allows a user to conduct a relatively broad search of large databases for relevant information that may not be explicitly stored in the databases. By

running statistical modules (claim 1) suggested by Simoudis et al. with module-specific parameters set by the user, for example, number of attributes selected, one of ordinary skill in the art may systematically reduce the number of attributes and respective values of said sample population (claim 1) by selecting the required attribute(s) by, for example, a number n (claim 4), a predetermined percentage (claim 5), or values exceeding a predetermined amount (claim 6), depending on requirements and extract a predictive model with more flexibility.

Accordingly, it would have been obvious to one having ordinary skill in the art at the time the invention was made to incorporate the entropy measure of Dash et al. to the statistical module of Simoudis et al. and perform data mining to samples having one or more desired attributes and respective values and all samples, then select attribute(s) to compute a model, because of the resulting efficiency and flexibility.

3.2 Regarding claims 13 and 25, Piatetsky-Shapiro teaches that the KID3 algorithm is extensible even to handle multifield conditions, such as, A1=a1 & A2=a2, in comparing attribute values for a target group of samples having the desired attribute value(s) to attribute values for all samples. See section 13.5. Therefore, patterns of full data set can be estimated by sample-derived rules with an estimated accuracy, see Abstract, i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. However, Piatetsky-Shapiro does not teach the selection of a subset of available attributes.

Simoudis et al. discloses a target data set constructed from a plurality of data sources, see col. 2, line 5 through line 26, and teaches the selection of a data analysis module, for example, a statistical module, to perform data mining to the selected data set. The user sets module-specific parameters. Once the user determined that the mining results are satisfactory based on the user's

queries or hypotheses, a predictive model is extracted based on such results. See col. 4, line 42 through line 57.

In order to generate a predictive model for a target group of samples having the desired attribute value(s) one of ordinary skill in the art would first be motivated by Piatetsky-Shapiro to compare attribute values for a target group of samples having the desired attribute value(s) to attribute values for all samples. Then to determine a difference between two data sets in a scientific and more efficient way one of ordinary skill in the art would be motivated by Simoudis et al. to run statistical modules to the above-mentioned data sets for data mining. Data mining allows a user to conduct a relatively broad search of large databases for relevant information that may not be explicitly stored in the databases. After running statistical modules with module-specific parameters set by the user, for example, number of attributes selected, one of ordinary skill in the art may systematically select the required attribute(s) depending on having a largest difference between the above-mentioned data sets and compute a predictive model with more flexibility.

Accordingly, it would have been obvious to one having ordinary skill in the art at the time the invention was made to select the statistical module suggested by Simoudis et al. and perform data mining to: (a) a target group of samples having the desired attribute value(s), and (b) all samples, then select attribute(s) to compute a model, because of the resulting efficiency and flexibility

3.3 Regarding claims 14-19, Piatetsky-Shapiro teaches comparing, see section 13.3, Rule-Interest Measures, attribute values for samples having a desired attribute value, see section 13.5, KID3 Algorithm, to attribute values for all samples, see section 13.4, Precomputing Field-Value

Statistics (claim 14). Therefore, patterns of full data set can be estimated by sample-derived rules with an estimated accuracy, see Abstract, i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. However, Piatetsky-Shapiro does not teach the selection of a subset of available attributes.

Simoudis et al. discloses a data mining system including a user interface, a target data set constructed from a plurality of data sources, and a server processor (claim 14), see col. 2, line 5 through line 26. Although Simoudis et al. does not mention a system memory, it is inherent for a processor with a system memory to execute a computer program. Simoudis et al. also teaches the selection of a data analysis module, for example, a statistical module, to perform data mining to the selected data set (claim 15). The user sets module-specific parameters. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results (claim 14). See col. 4, line 42 through line 57. However, Simoudis et al. does not teach the selection of a subset of available attribute based on entropy measure.

Dash et al. teaches an entropy measure (claim 16) for determining the relative importance of variables; see section 2. Dash et al. also discloses a simple way to decide how many variables should be kept for a task by choosing the first d variables if it is known that an application only needs d variables (claim 17). See section 3. Thus the user gains insight into the data after the important original features are known.

In order to set up a system to generate a predictive model for data having a desired attribute value one of ordinary skill in the art would first be motivated by Piatetsky-Shapiro to compare attribute values for samples having a desired attribute value to attribute values for all

samples (claim 14). Then to determine the statistical difference in a scientific and more efficient way one of ordinary skill in the art would be motivated by Dash et al. to apply an entropy measure (claim 16) to the above-mentioned data sets (claim 15) for data mining. Data mining allows a user to conduct a relatively broad search of large databases for relevant information that may not be explicitly stored in the databases. By setting up a system and running statistical modules (claim 15) suggested by Simoudis et al. with module-specific parameters set by the user, for example, number of attributes selected, one of ordinary skill in the art may systematically select the required attribute(s) (claim 14) by, for example, a predetermined number (claim 17), a predetermined percentage (claim 18), or values exceeding a predetermined amount (claim 19), depending on requirements and generate a predictive model (claim 14) with more flexibility.

Accordingly, it would have been obvious to one having ordinary skill in the art at the time the invention was made to set up a system and incorporate the entropy measure of Dash et al. to the statistical module of Simoudis et al. and perform data mining to samples having a desired attribute value and all samples, then select attribute(s) to compute a model, because of the resulting efficiency and flexibility.

3.4 Regarding claim 20, Piatetsky-Shapiro teaches that the KID3 algorithm is extensible even to handle multifield conditions, such as, A1=a1 & A2=a2, in comparing attribute values for a target subset of the plurality of samples having the desired attribute value(s) to attribute values for all of the samples. See sections 13.3-13.5. Therefore, patterns of full data set can be estimated by sample-derived rules with an estimated accuracy, see Abstract, i.e., a predictive

model of full data set can be generated by sample-derived rules with an estimated accuracy. However, Piatetsky-Shapiro does not teach the selection of a subset of available attributes.

Simoudis et al. discloses a data mining system including a user interface, a target data set constructed from a plurality of data sources, and a server processor, see col. 2, line 5 through line 26. Although Simoudis et al. does not mention a system memory, it is inherent for a processor with a system memory to execute a computer program. Simoudis et al. also teaches the selection of a data analysis module, for example, a statistical module, to perform data mining to the selected data set. The user sets module-specific parameters. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results. See col. 4, line 42 through line 57.

In order to set up a system to generate a predictive model for a target subset of samples having the desired attribute value(s) one of ordinary skill in the art would first be motivated by Piatetsky-Shapiro to compare attribute values for a target subset of samples having the desired attribute value(s) to attribute values for all samples. Then to determine a difference between two data sets in a scientific and more efficient way one of ordinary skill in the art would be motivated by Simoudis et al. to run statistical modules to the above-mentioned data sets for data mining. Data mining allows a user to conduct a relatively broad search of large databases for relevant information that may not be explicitly stored in the databases. After running statistical modules with module-specific parameters set by the user, for example, number of attributes selected, one of ordinary skill in the art may systematically select the required attribute(s) depending on having a largest difference between the above-mentioned data sets and compute a predictive model with more flexibility.

Accordingly, it would have been obvious to one having ordinary skill in the art at the time the invention was made to set up a system, select the statistical module suggested by Simoudis et al., and perform data mining to: (a) a target subset of samples having the desired attribute value(s), and (b) all samples, then select model attribute(s) and compute a model because of the resulting efficiency and flexibility.

3.5 Regarding claims 21-24, Piatetsky-Shapiro teaches comparing attribute values for samples having a desired attribute value to attribute values for all samples (claim 21). See sections 13.3-13.5. Therefore, patterns of full data set can be estimated by sample-derived rules with an estimated accuracy, see Abstract, i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy. However, Piatetsky-Shapiro does not teach the selection of a subset of available attributes.

Simoudis et al. discloses a target data set constructed from a plurality of data sources, see col. 2, line 5 through line 26, and teaches the selection of a data analysis module, for example, a statistical module, to perform data mining to the selected data set (claim 22). The user sets module-specific parameters. Once the user determined that the mining results are satisfactory based on the user's queries or hypotheses, a predictive model is extracted based on such results. See col. 4, line 42 through line 57. However, Simoudis et al. does not teach the selection of a subset of available attribute based on entropy measure.

Dash et al. teaches an entropy measure (claim 23) for determining the relative importance of variables; see section 2. Dash et al. also discloses a simple way to decide how many variables should be kept for a task by choosing the first d variables if it is known that an application only

needs d variables (claim 24). See section 3. Thus the user gains insight into the data after the important original features are known.

In order to generate a predictive model for data having a desired attribute value one of ordinary skill in the art would first be motivated by Piatetsky-Shapiro to compare attribute values for samples having a desired attribute value to attribute values for all samples (claim 21). Then to determine the statistical difference in a scientific and more efficient way one of ordinary skill in the art would be motivated by Dash et al. to apply an entropy measure (claim 23) to the above-mentioned data sets (claims 22 and 23) for data mining. Data mining allows a user to conduct a relatively broad search of large databases for relevant information that may not be explicitly stored in the databases. By running statistical modules (claims 22 and 23) suggested by Simoudis et al. with module-specific parameters set by the user, for example, number of attributes selected, one of ordinary skill in the art may systematically select the required attribute(s) (claim 21) by, for example, a number n (claim 24), depending on having a largest difference between the above-mentioned data sets and compute a predictive model with more flexibility.

Accordingly, it would have been obvious to one having ordinary skill in the art at the time the invention was made to incorporate the entropy measure of Dash et al. to the statistical module of Simoudis et al. and perform data mining to samples having a desired attribute value and all samples, then select attribute(s) to compute a model, because of the resulting efficiency and flexibility.

Applicant's Arguments

4. Applicants argue that the rejections made under 35 U.S.C. 103(a) are inapplicable because:

(1). Piatetsky-Shapiro only teaches an algorithm for discovering exact rules in databases. The claimed invention performs the comparison on attributes and respective values to obtain a target group (Reply, page 9, paragraph 3).

(2). The summary in Piatetsky-Shapiro does not preserve all of the field values and their relation to one another (Reply, page 10, last paragraph).

(3). Piatetsky-Shapiro does not teach comparing multiple attributes of a target population with attribute values of the entire sample population (Reply, page 11, paragraph 2).

(4). No citation has been provided in Piatetsky-Shapiro supporting the proposition that the full data set can be estimated by sample-derived rules and no reference to generating an estimated full data set is found in Piatetsky-Shapiro (Reply, page 11, paragraph 3).

(5). Combining Piatetsky-Shapiro, Dash et al., and Simoudis et al. does not teach the present invention because Piatetsky-Shapiro discovers exact rules and does not preserve the data (Reply, page 11, paragraph 4).

Response to Arguments

5. Applicant's arguments have been fully considered but they are not persuasive.

5.1 In response to Applicant's arguments (1)-(3), a recitation of the intended use of the claimed invention must result in a structural difference between the claimed invention and the

prior art in order to patentably distinguish the claimed invention from the prior art. If the prior art structure is capable of performing the intended use, then it meets the claim.

Piatetsky-Shapiro discloses KID3 Algorithm for discovery of exact rules. At the end, a cell for A=a contains the summary of all the file tuples satisfying A=a (page 235, line 34). The summary may preserve all of the field values and their relation to one another because what information has to be kept is determined by the type of the summary description we want to have at the end (page 236, lines 6-8). Therefore, one of ordinary skill in the art would be able to obtain a target group with one or more desired attributes and respective values, for example, A=a and B=b, by applying KID3 Algorithm and setting the summary to collect all the information of every sample which satisfying A=a and B=b within the sample population.

Piatetsky-Shapiro also discloses how to precompute field statistics (page 230, line 19) to get information such as "How many samples satisfy the condition of C=c?" (page 233, lines 13-16) in the sample population. With these statistics, the cardinality |C| of conditions C=c can be estimated efficiently and accurately (page 233, lines 19-20).

Besides, Piatetsky-Shapiro teaches that the rule-discovery task is to find K rules with the highest rule-interest function (page 231, line 28). Usually, the interest of rule A → B is computed as a function of p(A), the probability of A; p(B); p(A&B); or other parameters such as the domain sizes of A and B (page 231, lines23-26). The simplest function that Piatetsky-Shapiro mentioned is |A&B| - (|A||B|/N) (page 232, line 18).

The following example serves to explain what information and knowledge Piatetsky-Shapiro has disclosed. Let N be the total number of samples. A, B, C, D, and E are attributes of each sample. Target group is all the samples with A=a and B=b. To find out which attribute of

C, D, or E, or which attribute value of e1 or e2, is more sensitive for the purpose of generating a predictive model, one of ordinary skill in the art would be motivated by applying Piatetsky-Shapiro's algorithm for the following rules:

(A=a and B=b) → (C=c)

(A=a and B=b) → (D=d)

(A=a and B=b) → (E=e1) and (A=a and B=b) → (E=e2)

Therefore, |(A=a and B=b) & (C=c)|, |(A=a and B=b) & (D=d)|, |(A=a and B=b) & (E=e1)|, and |(A=a and B=b) & (E=e2)| are calculated. Statistics of |(A=a and B=b)|, |(C=c)|, |(D=d)|, |(E=e1)|, and |(E=e2)| are precomputed. By using the simplest rule-interest function a simple statistical measure will be calculated. Using a different rule-interest function, such as an entropy function, may generate a different statistical measure. Based on the calculated statistical measure, one of ordinary skill in the art will select the required attribute(s) or attribute value(s) easily for the purpose of generating a predictive model.

5.2 Regarding Applicant's argument (4), Piatetsky-Shapiro derives general formulas for estimating the accuracy of sample-derived rules on the full data set (page 229, lines 7-8). Rules can be used directly for providing high-level answers to user requests (page 230, line 13). Therefore, patterns of full data set can be estimated by sample-derived rules with an estimated accuracy, i.e., a predictive model of full data set can be generated by sample-derived rules with an estimated accuracy.

5.3 Regarding Applicant's argument (5), Piatetsky-Shapiro not only discovers exact rules, the KID3 algorithm also may preserve the data as explained in section **5.1** above. Therefore,

Examiner believes that all the claimed inventions have been disclosed by combining Piatetsky-Shapiro, Dash et al., and Simoudis et al.

Conclusion

6. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

7. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

8. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Herng-der Day whose telephone number is (703) 305-5269. The examiner can normally be reached on 8:30 - 17:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Kevin J Teska can be reached on (703) 305-9704. The fax phone numbers for the organization where this application or proceeding is assigned are (703) 746-7239 for regular communications and (703) 746-7238 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be direcied to the receptionist whose telephone number is (703) 305-3900.

Herng-der Day
July 11, 2002

Samuel Broda

SAMUEL BRODA, ESQ.
PATENT EXAMINER